

RECEIVED

11

APR 09 2003

TECH CENTER



1600

RAW SEQUENCE LISTING

DATE: 04/04/2003

PATENT APPLICATION: US/09/770,875A

TIME: 14:45:53

Input Set : A:\20264251.app

Output Set: N:\CRF4\04042003\I770875A.raw

3 <110> APPLICANT: CIEPLAK, WITOLD
 5 <120> TITLE OF INVENTION: PERTUSSIN TOXIN GENE: CLONING AND EXPRESSION OF
 6 . PROTECTIVE ANTIGEN
 8 <130> FILE REFERENCE: 2026-4253US7c
 10 <140> CURRENT APPLICATION NUMBER: 09/770,875A
 11 <141> CURRENT FILING DATE: 2001-01-26
 13 <150> PRIOR APPLICATION NUMBER: 07/311,612
 14 <151> PRIOR FILING DATE: 1989-02-15
 16 <150> PRIOR APPLICATION NUMBER: 07/542,149
 17 <151> PRIOR FILING DATE: 1990-06-22
 19 <150> PRIOR APPLICATION NUMBER: 08/483,326
 20 <151> PRIOR FILING DATE: 1995-06-07
 22 <150> PRIOR APPLICATION NUMBER: 09/128,911
 23 <151> PRIOR FILING DATE: 1998-08-04
 25 <160> NUMBER OF SEQ ID NOS: 28
 27 <170> SOFTWARE: PatentIn Ver. 2.1
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 184
 31 <212> TYPE: DNA
 32 <213> ORGANISM: Bordetella pertussis
 34 <400> SEQUENCE: 1
 35 cccgggacag ggcggcgccc ggcggtcgcg ggtccgcgcc ctggcggtggt tcctgccatc 60
 36 cggcgcgatg acgcattctt ccccgccctt ggccgacgtt ccttatgtgc tggatgaagac 120
 37 caatatgggtg gtcaccagcg tagccatgaa gccgtatgaa gtcaccccca cgcggtatgct 180
 38 gggtc 184
 41 <210> SEQ ID NO: 2
 42 <211> LENGTH: 61
 43 <212> TYPE: PRT
 44 <213> ORGANISM: Bordetella pertussis
 46 <400> SEQUENCE: 2
 47 Pro Gly Gln Gly Gly Ala Arg Arg Ser Arg Val Arg Ala Leu Ala Trp
 48 1 5 10 15
 50 Leu Leu Ala Ser Gly Ala Met Thr His Leu Ser Pro Ala Leu Ala Asp
 51 20 25 30
 53 Val Pro Tyr Val Leu Val Lys Thr Asn Met Val Val Thr Ser Val Ala
 54 35 40 45
 56 Met Lys Pro Tyr Glu Val Thr Pro Thr Arg Met Leu Val
 57 50 55 60
 60 <210> SEQ ID NO: 3
 61 <211> LENGTH: 17
 62 <212> TYPE: DNA
 63 <213> ORGANISM: Bordetella pertussis
 65 <220> FEATURE:

ENTERED

RAW SEQUENCE LISTING

DATE: 04/04/2003

PATENT APPLICATION: US/09/770,875A

TIME: 14:45:53

Input Set : A:\20264251.app

Output Set: N:\CRF4\04042003\I770875A.raw

66 <223> OTHER INFORMATION: Purine (P) R=G or A; Y=T or C; N=A, C, G, or T

68 <220> FEATURE:

69 <221> NAME/KEY: modified_base

70 <222> LOCATION: (6)

72 <220> FEATURE:

73 <221> NAME/KEY: modified_base

74 <222> LOCATION: (9)

76 <220> FEATURE:

77 <221> NAME/KEY: modified_base

78 <222> LOCATION: (12)

80 <220> FEATURE:

81 <221> NAME/KEY: modified_base

82 <222> LOCATION: (15)

84 <400> SEQUENCE: 3

W--> 85 atgaarccnt aygargt

17

88 <210> "SEQ ID NO": 4

89 <211> LENGTH: 30

90 <212> TYPE: PRT

91 <213> ORGANISM: Bordetella pertussis

93 <220> FEATURE:

94 <223> OTHER INFORMATION: Xaa = Any amino acid; the 8th Val and 4th Pro are
95 questionable.

97 <220> FEATURE:

98 <221> NAME/KEY: VARIANT

99 <222> LOCATION: (15)

101 <400> SEQUENCE: 4

W--> 102 Asp Val Pro Tyr Val Leu Val Lys Thr Asn Met Val Val Thr Xaa Val

103 1 5 10 15

105 Ala Met Lys Pro Tyr Glu Val Val Pro Pro Arg Met Leu Val

106 20 25 30

109 <210> SEQ ID NO: 5

110 <211> LENGTH: 4210

111 <212> TYPE: DNA

112 <213> ORGANISM: Bordetella pertussis

114 <220> FEATURE:

115 <221> NAME/KEY: CDS

116 <222> LOCATION: (609)..(1310)

118 <220> FEATURE:

119 <221> NAME/KEY: CDS

120 <222> LOCATION: (1434)..(2030)

122 <220> FEATURE:

123 <221> NAME/KEY: CDS

124 <222> LOCATION: (2153)..(2482)

126 <220> FEATURE:

127 <221> NAME/KEY: CDS

128 <222> LOCATION: (2557)..(2856)

130 <220> FEATURE:

131 <221> NAME/KEY: CDS

132 <222> LOCATION: (3026)..(3622)

RAW SEQUENCE LISTING

DATE: 04/04/2003

PATENT APPLICATION: US/09/770,875A

TIME: 14:45:53

Input Set : A:\20264251.app

Output Set: N:\CRF4\04042003\I770875A.raw

134 <400> SEQUENCE: 5

```

135 gaattcgtcg cctcgccctg gttcgccgctc atggcccccagggaaccga cccaagata 60
137 atcgtcctgc tcaaccgcca catcaacgag gcgctgcagt ccaaggcggt cgctcgaggcc 120
139 ttgcccggcc aaggcgccac gccgggtcatc gccacgcccgc atcagacccg cggcttcac 180
141 gcagacgaga tccagcgctg ggccggcgctc gtgcgcgaaa ccggcgccaa gctgaagtag 240
143 cagcgagacc ctccaacgcg ccatccccgt ccggccggca ccatcccgca tacgtgttg 300
145 caaccgcaa cgcgcatgcg tgcagattcg tcgtacaaaa ccctcgattc ttccgtacat 360
147 cccgctactg caatccaaca cggcatgaac gtccttcgg cgcaaagtcg cgcgatggt 420
149 ccggtcaccg tccggaccgt gctgaccccc ctgccatggt gtgatcccta aaataggcac 480
151 catcaaaacg cagaggggaa gacgggatgc gttgcactcg ggcaattcg caaaccgcaa 540
153 gaacaggctg gctgacgtgg ctggcgattc ttgccgtcac ggccggcggtg acttcggccg 600
155 catgggcc gac gat cct ccc gcc acc gta tac cgc tat gac tcc cgc ccg 650
156 Asp Asp Pro Pro Ala Thr Val Tyr Arg Tyr Asp Ser Arg Pro
157 1 5 10
159 ccg gag gac gtt ttc cag aac gga ttc acg gcg tgg gga aac aac gac 698
160 Pro Glu Asp Val Phe Gln Asn Gly Phe Thr Ala Trp Gly Asn Asn Asp
161 15 20 25 30
163 aat gtg ctc gac cat ctg acc gga cgt tcc tgc cag gtc ggc agc agc 746
164 Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys Gln Val Gly Ser Ser
165 35 40 45
167 aac agc gct ttc gtc tcc acc agc agc agc cgg cgc tat acc gag gtc 794
168 Asn Ser Ala Phe Val Ser Thr Ser Ser Ser Arg Arg Tyr Thr Glu Val
169 50 55 60
171 tat ctc gaa cat cgc atg cag gaa gcg gtc gag gcc gaa cgc gcc ggc 842
172 Tyr Leu Glu His Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly
173 65 70 75
175 agg ggc acc ggc cac ttc atc ggc tac atc tac gaa gtc cgc gcc gac 890
176 Arg Gly Thr Gly His Phe Ile Gly Tyr Ile Tyr Glu Val Arg Ala Asp
177 80 85 90
179 aac aat ttc tac ggc gcc gcc agc tcg tac ttc gaa tac gtc gac act 938
180 Asn Asn Phe Tyr Gly Ala Ala Ser Ser Tyr Phe Glu Tyr Val Asp Thr
181 95 100 105 110
183 tat ggc gac aat gcc ggc cgt atc ctc gcc ggc gcg ctg gcc acc tac 986
184 Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly Ala Leu Ala Thr Tyr
185 115 120 125
187 cag agc gaa tat ctg gca cac cgg cgc att ccg ccc gaa aac atc cgc 1034
188 Gln Ser Glu Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile Arg
189 130 135 140
191 agg gta acg cgg gtc tat cac aac ggc atc acc ggc gag acc acg acc 1082
192 Arg Val Thr Arg Val Tyr His Asn Gly Ile Thr Gly Glu Thr Thr Thr
193 145 150 155
195 acg gag tat tcc aac gct cgc tac gtc agc cag cat act cgc gcc aat 1130
196 Thr Glu Tyr Ser Asn Ala Arg Tyr Val Ser Gln His Thr Arg Ala Asn
197 160 165 170
199 ccc aac ccc tac aca tcg cga agg tcc gta gcg tcg atc gtc ggc aca 1178
200 Pro Asn Pro Tyr Thr Ser Arg Arg Ser Val Ala Ser Ile Val Gly Thr
201 175 180 185 190
203 ttg gtg cgc atg gcg ccg gtg ata ggc gct tgc atg gcg cgg cag gcc 1226
204 Leu Val Arg Met Ala Pro Val Ile Gly Ala Cys Met Ala Arg Gln Ala

```

RAW SEQUENCE LISTING

DATE: 04/04/2003

PATENT APPLICATION: US/09/770,875A

TIME: 14:45:53

Input Set : A:\20264251.app

Output Set: N:\CRF4\04042003\I770875A.raw

| | | | | |
|-----|--|------|-----|-----|
| 205 | 195 | 200 | 205 | |
| 207 | gaa agc tcc gag gcc atg gca gcc tgg tcc gaa cgc gcc ggc gag gcg | 1274 | | |
| 208 | Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu Ala | | | |
| 209 | 210 | 215 | 220 | |
| 211 | atg gtt ctc gtg tac tac gaa agc atc gcg tat tcg ttctagacct | 1320 | | |
| 212 | Met Val Leu Val Tyr Tyr Glu Ser Ile Ala Tyr Ser | | | |
| 213 | 225 | 230 | | |
| 215 | ggcccagccc cgcccaactc cggttaattca acagcatgcc gatcgaccgc aagacgctct | 1380 | | |
| 217 | gccatctcct gtccgttctg ccgttgcccc tctcggatc tcacgtggcg cgg gcc | 1436 | | |
| 218 | | | Ala | |
| 219 | | | 235 | |
| 221 | tcc acg cca ggc atc gtc att ccg ccg cag gaa cag att acc cag cat | 1484 | | |
| 222 | Ser Thr Pro Gly Ile Val Ile Pro Pro Gln Glu Gln Ile Thr Gln His | | | |
| 223 | 240 | 245 | 250 | |
| 225 | ggc agc ccc tat gga cgc tgc gcg aac aag acc cgt gcc ctg acc gtg | 1532 | | |
| 226 | Gly Ser Pro Tyr Gly Arg Cys Ala Asn Lys Thr Arg Ala Leu Thr Val | | | |
| 227 | 255 | 260 | 265 | |
| 229 | gcg gaa ttg cgc ggc agc ggc gat ctg cag gag tac ctg cgt cat gtg | 1580 | | |
| 230 | Ala Glu Leu Arg Gly Ser Gly Asp Leu Gln Glu Tyr Leu Arg His Val | | | |
| 231 | 270 | 275 | 280 | |
| 233 | acg cgc ggc tgg tca ata ttt gcg ctc tac gat ggc acc tat ctc ggc | 1628 | | |
| 234 | Thr Arg Gly Trp Ser Ile Phe Ala Leu Tyr Asp Gly Thr Tyr Leu Gly | | | |
| 235 | 285 | 290 | 295 | |
| 237 | ggc gaa tat ggc ggc gtg atc aag gac gga aca ccc ggc ggc gca ttc | 1676 | | |
| 238 | Gly Glu Tyr Gly Gly Val Ile Lys Asp Gly Thr Pro Gly Gly Ala Phe | | | |
| 239 | 300 | 305 | 310 | 315 |
| 241 | gac ctg aaa acg acg ttc tgc atc atg acc acg cgc aat acg ggt caa | 1724 | | |
| 242 | Asp Leu Lys Thr Thr Phe Cys Ile Met Thr Thr Arg Asn Thr Gly Gln | | | |
| 243 | 320 | 325 | 330 | |
| 245 | ccc gca acg gat cac tac tac agc aac gtc acc gcc act cgc ctg ctc | 1772 | | |
| 246 | Pro Ala Thr Asp His Tyr Tyr Ser Asn Val Thr Ala Thr Arg Leu Leu | | | |
| 247 | 335 | 340 | 345 | |
| 249 | tcc agc acc aac agc agg cta tgc gcg gtc ttc gtc aga agc ggg caa | 1820 | | |
| 250 | Ser Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val Arg Ser Gly Gln | | | |
| 251 | 350 | 355 | 360 | |
| 253 | ccg gtc att ggc gcc tgc acc agc ccg tat gac ggc aag tac tgg agc | 1868 | | |
| 254 | Pro Val Ile Gly Ala Cys Thr Ser Pro Tyr Asp Gly Lys Tyr Trp Ser | | | |
| 255 | 365 | 370 | 375 | |
| 257 | atg tac agc cgg ctg cgg aaa atg ctt tac ctg atc tac gtg gcc ggc | 1916 | | |
| 258 | Met Tyr Ser Arg Leu Arg Lys Met Leu Tyr Leu Ile Tyr Val Ala Gly | | | |
| 259 | 380 | 385 | 390 | 395 |
| 261 | atc tcc gta cgc gtc cat gtc agc aag gaa cag tat tac gac tat | 1964 | | |
| 262 | Ile Ser Val Arg Val His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr | | | |
| 263 | 400 | 405 | 410 | |
| 265 | gag gac gca acg ttc gag act tac gcc ctt acc ggc atc tcc atc tgc | 2012 | | |
| 266 | Glu Asp Ala Thr Phe Glu Thr Tyr Ala Leu Thr Gly Ile Ser Ile Cys | | | |
| 267 | 415 | 420 | 425 | |
| 269 | aat cct gga tca tcc tta tgctgagacg cttccccact cgaaccaccg | 2060 | | |
| 270 | Asn Pro Gly Ser Ser Leu | | | |

RAW SEQUENCE LISTING

DATE: 04/04/2003

PATENT APPLICATION: US/09/770,875A

TIME: 14:45:53

Input Set : A:\20264251.app

Output Set: N:\CRF4\04042003\I770875A.raw

```

271          430
273 ccccgaggaca gggcggcgcc cggcggtcgc gcatgcgcgc cctggcggtgg ttgctggcat 2120
275 ccggcgcgat gacgcattctt tcccccgccc tg gcc gac gtt cct tat gtg ctg 2173
276                               Ala Asp Val Pro Tyr Val Leu
277                               435                               440
279 gtg aag acc aat atg gtg gtc acc agc gta gcc atg aag ccg tat gaa 2221
280 Val Lys Thr Asn Met Val Val Thr Ser Val Ala Met Lys Pro Tyr Glu
281                               445                               450                               455
283 gtc acc ccg acg cgc atg ctg gtc tgc ggc atc gcc gcc aaa ctg ggc 2269
284 Val Thr Pro Thr Arg Met Leu Val Cys Gly Ile Ala Ala Lys Leu Gly
285                               460                               465                               470
287 gcc gcg gcc agc agc ccg gac gcg cac gtg ccg ttc tgc ttc ggc aag 2317
288 Ala Ala Ala Ser Ser Pro Asp Ala His Val Pro Phe Cys Phe Gly Lys
289                               475                               480                               485
291 gat ctc aag cgt ccc ggc agc agt ccc atg gaa gtc atg ttg cgc gcc 2365
292 Asp Leu Lys Arg Pro Gly Ser Ser Pro Met Glu Val Met Leu Arg Ala
293                               490                               495                               500
295 gtc ttc atg caa caa cgg ccg ctg cgc atg ttt ctg ggt ccc aag caa 2413
296 Val Phe Met Gln Gln Arg Pro Leu Arg Met Phe Leu Gly Pro Lys Gln
297 505                               510                               515                               520
299 ctc act ttc gaa ggc aag ccc gcg ctc gaa ctg atc cgg atg gtc gaa 2461
300 Leu Thr Phe Glu Gly Lys Pro Ala Leu Glu Leu Ile Arg Met Val Glu
301                               525                               530                               535
303 tgc agc ggc aag cag gat tgc ccctgaaggc gaaccccatg cataccatcg 2512
304 Cys Ser Gly Lys Gln Asp Cys
305                               540
307 catccatcct gttgtccgtg ctccggcatat acagcccggc tgac gtc gcc ggc ttg 2568
308                               Val Ala Gly Leu
309                               545
311 ccg acc cat ctg tac aag aac ttc act gtc cag gag ctg gcc ttg aaa 2616
312 Pro Thr His Leu Tyr Lys Asn Phe Thr Val Gln Glu Leu Ala Leu Lys
313                               550                               555                               560
315 ctg aag ggc aag aat cag gag ttc tgc ctg acc gcc ttc atg tcg ggc 2664
316 Leu Lys Gly Lys Asn Gln Glu Phe Cys Leu Thr Ala Phe Met Ser Gly
317                               565                               570                               575
319 aga agc ctg gtc cgg gcg tgc ctg tcc gac gcg gga cac gag cac gac 2712
320 Arg Ser Leu Val Arg Ala Cys Leu Ser Asp Ala Gly His Glu His Asp
321 580                               585                               590                               595
323 acg tgg ttc gac acc atg ctt ggc ttt gcc ata tcc gcg tat gcg ctc 2760
324 Thr Trp Phe Asp Thr Met Leu Gly Phe Ala Ile Ser Ala Tyr Ala Leu
325                               600                               605                               610
327 aag agc cgg atc gcg ctg acg gtg gaa gac tcg ccg tat ccg ggc act 2808
328 Lys Ser Arg Ile Ala Leu Thr Val Glu Asp Ser Pro Tyr Pro Gly Thr
329                               615                               620                               625
331 ccc ggc gat ctg ctc gaa ctg cag atc tgc ccg ctc aac gga tat tgc 2856
332 Pro Gly Asp Leu Leu Glu Leu Gln Ile Cys Pro Leu Asn Gly Tyr Cys
333                               630                               635                               640
335 gaatgaaccc ttccggagggt ttccgacgttt ccgcgcaatc cgcttgagac gatcttccgc 2916
337 cctggttcca ttccgggaac accgcaacat gctgatcaac aacaagaagc tgcttcatca 2976

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/770,875A

DATE: 04/04/2003
TIME: 14:45:54

Input Set : A:\20264251.app
Output Set: N:\CRF4\04042003\I770875A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 9
Seq#:4; Xaa Pos. 15